2006-760: AN ANALYSIS OF GRADUATION RATES AT RESEARCH UNIVERSITIES

Cindy Veenstra, University of Michigan
Cindy Veenstra is a Ph.D. Candidate in Industrial and Operations Engineering at the University of Michigan. She is a quality consultant and has an interest in applying Quality Engineering techniques to Engineering Education research. Her professional affiliations include ASEE, ASQ and INFORMS. She may be reached at cpveenst@umich.edu.

Gary D. Herrin, University of Michigan
Gary D. Herrin, Ph.D. is Interim Associate Dean of Undergraduate Education of the College of Engineering and Professor of Industrial and Operations Engineering at the University of Michigan. He has authored over 150 papers in Industrial Engineering. His professional affiliations include ASEE and ASQ. He may be reached at gdherrin@umich.edu

An Analysis of Graduation Rates at Research Universities Using the Education Trust Database

Abstract

Many large research universities (i.e. with doctoral programs) have high undergraduate retention rates as measured by their 6-year graduation rates. The variation in graduation rates between schools and the relationship between graduation rates and student preparation levels were examined for these universities. A university was defined as successful with respect to its graduation rate if it had a 6-year graduation rate of 75% of more. With a preference for universities with strong technical programs such as engineering, universities were further selected with a high percentage (25% or more) of students pursuing a STEM major. The Education Trust database system was used to collect these data. Data of particular interest were the overall graduation rates, Under-Represented Minority (URM) graduation rates and the difference in graduation rates between the overall student population and the URM student population. In this research, a high priority was given to assessing whether there was a strong relationship between this difference in graduation rates and the preparation level of the student body as measured by the median SAT and the amount of money that flowed into each program (denoted as student related expenditures in the Education Trust database).

A strong relationship between the differences in graduation rates was observed using multiple regression analysis. A significant difference in this relationship was found between private and public universities. In addition, Data Envelopment Analysis was used to determine which universities were "benchmark" universities or universities that were best in class relative to having a minimal gap between the six year graduation rates of the overall student population and the URM student population. Data Envelopment Analysis is a linear programming technique that establishes an "efficiency frontier" subset of points with maximum output values relative to the input values.

This presentation will demonstrate the techniques that can be used by other researchers to benchmark a group of universities relative to graduation rates. If graduation rates of colleges of engineering were available, the same techniques could be applied. With the use of the Data Envelopment Analysis, an engineering university can compare itself to an appropriate benchmark university and use this information to improve its processes for achieving a higher retention rate of all students.

The Need for a National Database on Graduation Rates for Engineering Colleges

This research supports the need for a database to define the relationship between the graduation rates of engineering colleges and college preparation level as measured by the SAT or ACT scores. For continuous improvement there is a need to benchmark engineering colleges by graduation rates. There is no recognized national database on graduation rates of engineering colleges. Discussions with ASEE support staff indicated that ASEE does not request graduation rates in their college database. Other efforts at identifying a national engineering college database also were futile.

The most prominent university level aggregate database found was the Education Trust Database. This paper explores relationships observable between graduation rates and other relevant variables, such as college preparedness, using the Education Trust Database...

The analysis includes large research universities with a high percent of STEM majors. A limitation of this analysis is that, for large research universities, the results represent experience of the entire university not just the engineering college or STEM disciplines. However, the research documented in this paper shows the benefits of such an analysis and demonstrates the potential of a national engineering college database, if it existed.

Introduction

The research literature is robust with articles on retention rates (graduation rates) of universities. Research indicates that one of the major predictors of graduation rate is academic preparedness upon entering college. [5]

Much emphasis in academia has been placed on the enrollment of minority students with much less emphasis on the graduation rate of minorities. For under-represented minorities the 6-year graduation rate for all four-year colleges and universities is only 45%, compared to the overall 6-year graduation rate of 55%. [1] Among the large research universities that grant doctorates, the 6-year graduation rate for under-represented minorities is 58% compared to approximately a 67% 6-year graduation rate overall. (These statistics for large research universities are based on an average across all universities that are classified as Doctoral research extensive in the Education Trust 2003 database.)

Yet, there are success stories. Some universities have higher graduation rates of 85% or more for both the overall student population and the minority student population. By identifying universities that have only a small gap in graduation rates between minorities and the overall student population, benchmark universities can be established. Success stories from these benchmark universities may give other universities the process tools to improve their graduation rates.

Research Objectives

This research looks at the variation in graduation rates and graduation rate gap between minorities and the overall student population for large research universities that have been successful with graduation rates. Variation among universities and over time is explored. Because of an interest in technical degree-seekers, only universities with a high percent of degree seekers in science, technology, engineering and math (STEM) were selected.

Graduation rate gap is defined as the difference of the 6-year graduation rate of under-represented minorities (URM) from the graduation rate of all students. Relationships between variables available in the database were explored with a stepwise regression. Benchmarking of successful universities was accomplished using Data Envelopment Analysis.

Database Source

The College Results Online database at the Education Trust website[4] was used. Education Trust is a non-profit organization that developed this database using data from the U.S. Department of Education's National Center for Education Statistics Graduation Rate Survey (GRS). The data collection system is known as Integrated Post-secondary Education Data System (IPEDS). The 6-year graduation rates for the year 2003 will be used.

The College results Online Database has several compelling characteristics. With an Excel format, it is easy to use and transform to a statistical analysis. With the many filters available, one can filter on universities with a higher technical emphasis through the variable called Percent of STEM majors. Further, it has a large number of variables including the educational institutions' median ACT and SAT sub-scores, graduation rates from 1997 to 2003 and first year retention rates.

In this analysis, the Education Trust College Results Online Database was used with the following filters:

- Universities that are classified as Doctoral Research Extensive using the Carnegie classification [2]
- Universities with a 6-year overall graduation rate of more than 75% (very successful)
- Universities with an STEM percent greater than 25% (a high percent of technical degree seekers)

Twenty-five universities were identified; the data analysis included 24 universities. The California Institute of Technology was considered an outlier with its high SAT median and small student population. The list of universities included in this study is in Table 1.

Table 1: Universities Included in Study

University Of Wisconsin-Madison	Rice University
University Of Illinois at Urbana	University Of California-Irvine
Lehigh University	Cornell University
University Of Michigan-Ann Arbor	University Of Rochester
University Of California-Davis	Johns Hopkins University
Tufts University	Stanford University
University Of California-Berkeley	Princeton University
University Of California-San Diego	Duke University
Massachusetts Institute of Technology	Columbia University
Rensselaer Polytechnic Institute	Vanderbilt University
Texas A & M University	Case Western Reserve University
Carnegie Mellon University	Washington University in St Louis

Nomenclature

Definitions of variables used in this study are presented as follows:

1. 6-year graduation rate - rate of graduation of students who started 6 years prior as freshmen. This would include students who graduated in 4 or 5 years.

2. Graduation Gap- the difference between the overall student 6-year graduation rate and the under-represented minority (URM) 6-year graduation rate.

3. Under-Represented Minority (URM) – students who are identified in ethnicity as African-American, Latinos, and Native American

4. Doctoral Research Extensive Universities- This is a classification defined by the Carnegie Foundation and used in the Education Trust Database system. A doctoral research extensive university has extensive bachelor programs and graduate school programs and awards at least 50 doctoral degrees per year in at least 15 disciplines.

5. STEM Percent- Percent of Students who have a major in Science, Technology, Engineering, or Mathematics.

6. Median SAT – the average of the 25th percentile and 75th percentile SAT scores. Documentation is available at the Education Trust website.

7. Student related expenditures/FTE – This is the student related expenditures per full time equivalent student. A full time student is defined as a student who is taking at least 12 credit hours of courses. The student related expenditures is calculated as the sum of the instruction expenditures plus the student services expenditures plus the proportion of academic support expenditures related to instruction of students. Academic support overhead for public service and research is not included. Academic support for tutoring and mentoring would be included. This formula was developed by the National Center for Higher Education Management Systems (NCHEMS).

Assumptions

The 2003 College Results Online Database includes the 6-year graduation rates for all students who started as freshman in 1997. (with certain exclusions defined by the Department of Education, documented on the Education Trust website) The database also includes the Median SAT from in the 2003 database. This is the median SAT for the 2002-2003 freshman class. When regressions of the graduation rates versus the median SAT are presented, it is assumed that the educational institution's median SAT is stable over time; i.e., the median SAT for 2002-2003 is close to the median SAT for the 1997-1998 freshman class. The same is true for the predictor %URM and student related expenditures/FTE included in this analysis.

Major Findings

- For this elite subset of universities with high graduation rates, the average graduation rate has shifted in the positive direction from 82.2% to 85.5% with a slight decrease in variation.

- The 5-year and 6-year graduation rates are more stable than the 4-year graduation rates.

- Using regression analysis, there is a statistically significant relationship between the overall 6-year graduation rate and median SAT.

- Using regression analysis, there is a statistically significant relationship between the overall 6-year graduation rate and the percentage of students who are under-represented minorities

- Using regression analysis, there is a statistically significant relationship between the graduation rate gap and student related expenditures/FTE.

- Using data envelopment analysis, the three large research universities that serve as benchmarks for similar universities are: Texas A&M University, Case Western Reserve University, and Washington University in St. Louis.

Variation Study of Graduation Rates

In any variation analysis, the stability of the statistic over time is an important consideration. How stable is the graduation rate statistic over time? For a given university, the graduation rate varies only a small percentage year to year. In a 6-year trend analysis of 6-year graduation rates from 1998 to 2003, the average improvement for all 24 universities was 3.2% with a 95% confidence interval on the average of 2.1 to 4.3%. Figure 1 shows the histogram for the six-year differences. One university had an improvement as large as 10%.

Figure 2 shows the distribution of the 2003 graduation rates compared to the 1998 graduation rates. Both distributions pass a statistical test for normality. With the shift in distributions and a smaller standard deviation in 2003, an improvement in graduation rates over six years is visible in Figure 2. This is positive evidence that the large research universities are progressing towards higher retention rates.
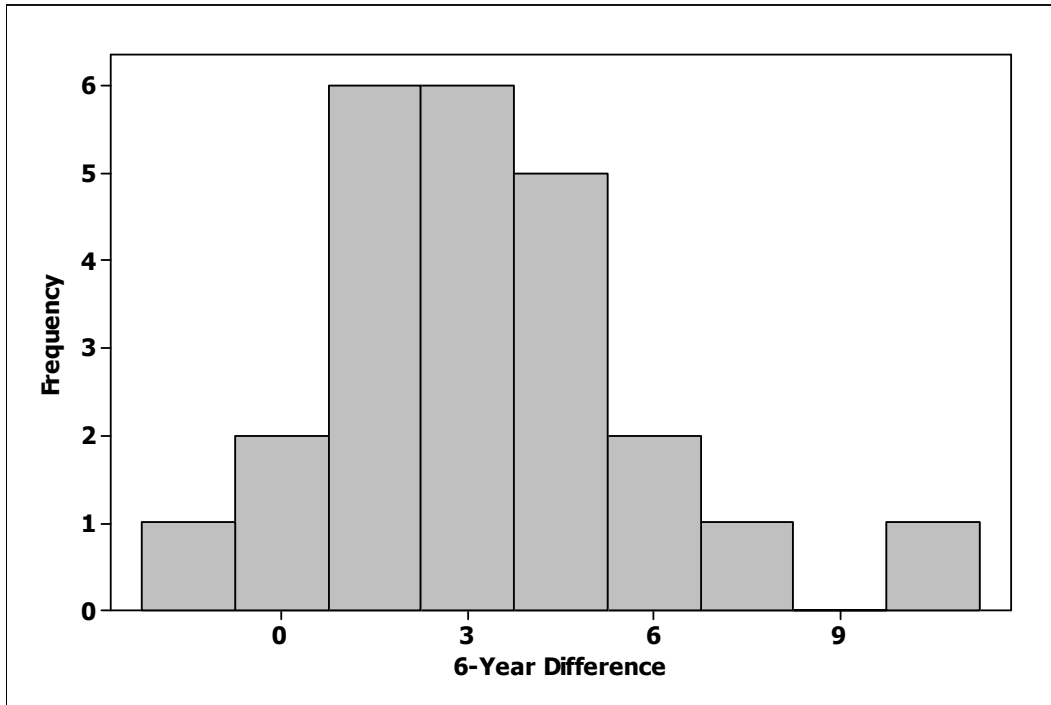
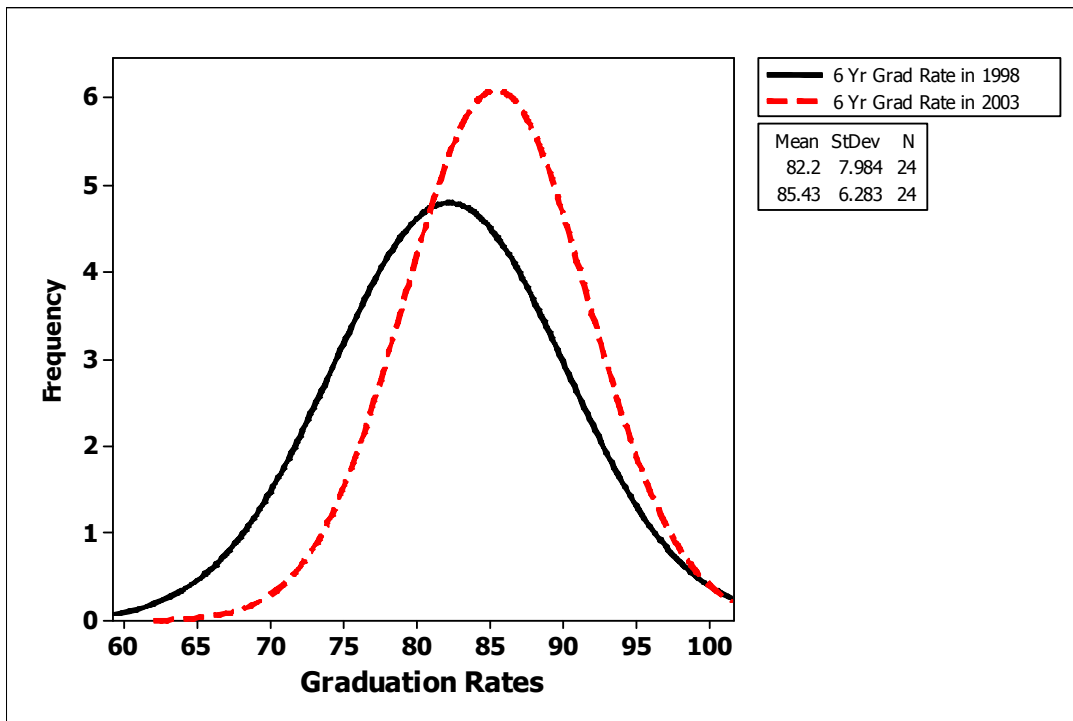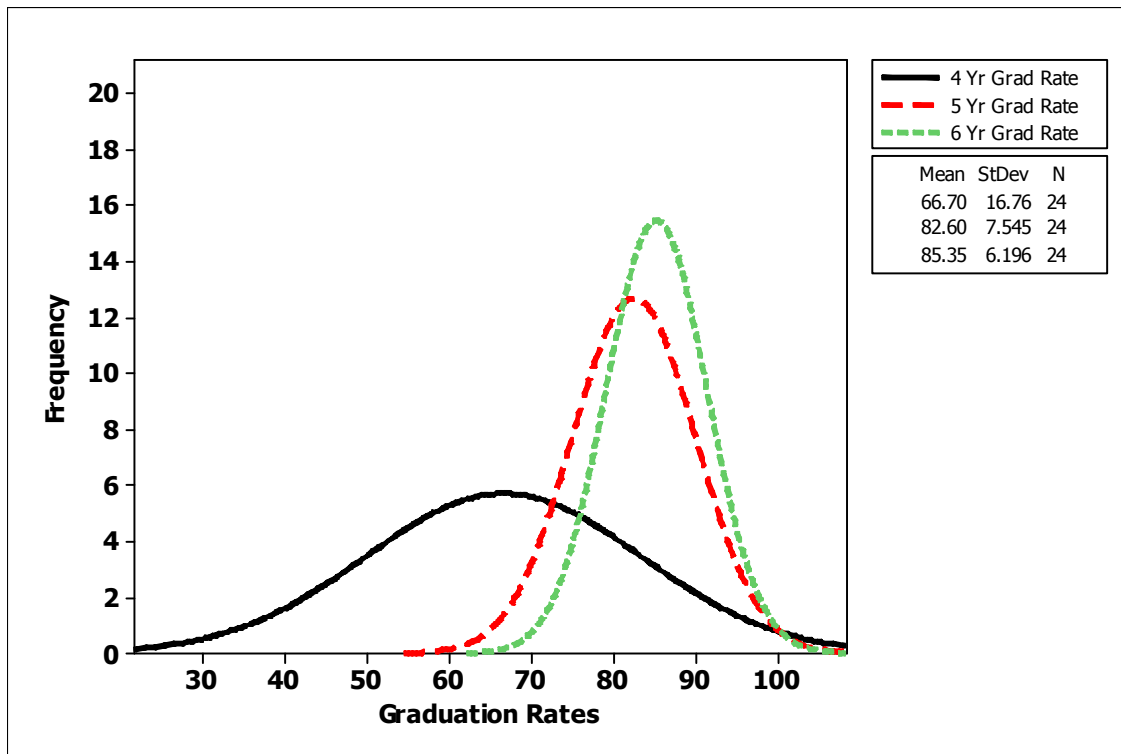**Figure 1: Histogram of 6-Year Differences in Graduation Rates**

Figure 2: Fitted Distribution of Graduation Rates in 2003 Compared to the Fitted Distribution of Graduation Rates in 1998

Another consideration in graduation rates is to how much difference exists between the 4-year graduation rates and the 6-year graduation rates. For a variety of reasons, some students require more than 4 years to graduate. The College Results Online database records the 4-year, 5-year and 6-year graduation rates. For the freshmen class of 1997, these graduation rates are shown in Figure 3. Notice the large difference in distributions between the 4-year and 5-year graduation rates both in the average graduation rate and the standard deviation of the distributions.


	4 Yr Grad Rate	5 Yr Grad Rate	6 Yr Grad Rate

Mean	StDev	N
66.70	16.76	24
82.60	7.545	24
85.35	6.196	24

Figure 3: Comparison of the Fitted Distributions of the 4-, 5-, and 6- year Graduation Rates

**Exploratory Data Analysis with Regression Analysis**

The median SAT score for a university is an overall measure of the academic preparedness of the student body of a university. From the research literature, it is expected that there will be a strong relationship between the overall graduation rate (of the freshman class six years prior) and the median SAT. As an added dimension, the universities are identified as either public or private educational institutions. The solid circles represent the private universities and the open circles represent the public universities. The dashed line is the 95% Prediction Interval associated with the regression. Figure 4 shows the linear relationship between the 6-year graduation rate and the median SAT.
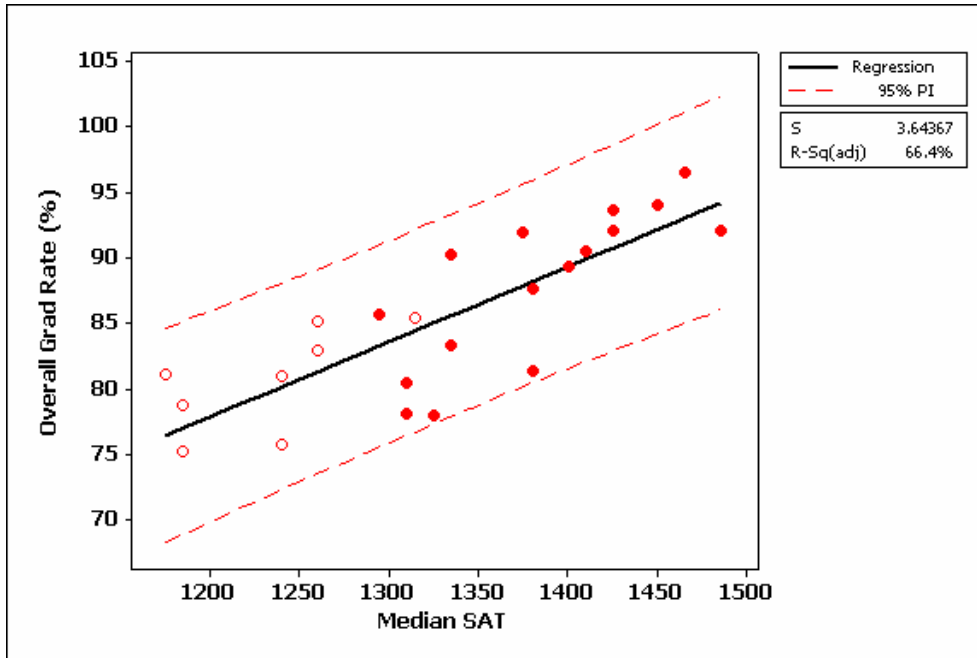
Figure 4: 6-year Overall Graduation Rate versus Median SAT
(Solid circles denote private universities; open circles denote public universities)

In addition, a stepwise linear regression showed a strong relationship between the 6-year overall graduation rates and the percent of the student body that was considered under-represented minorities (% URM). This is shown in Figure 5.
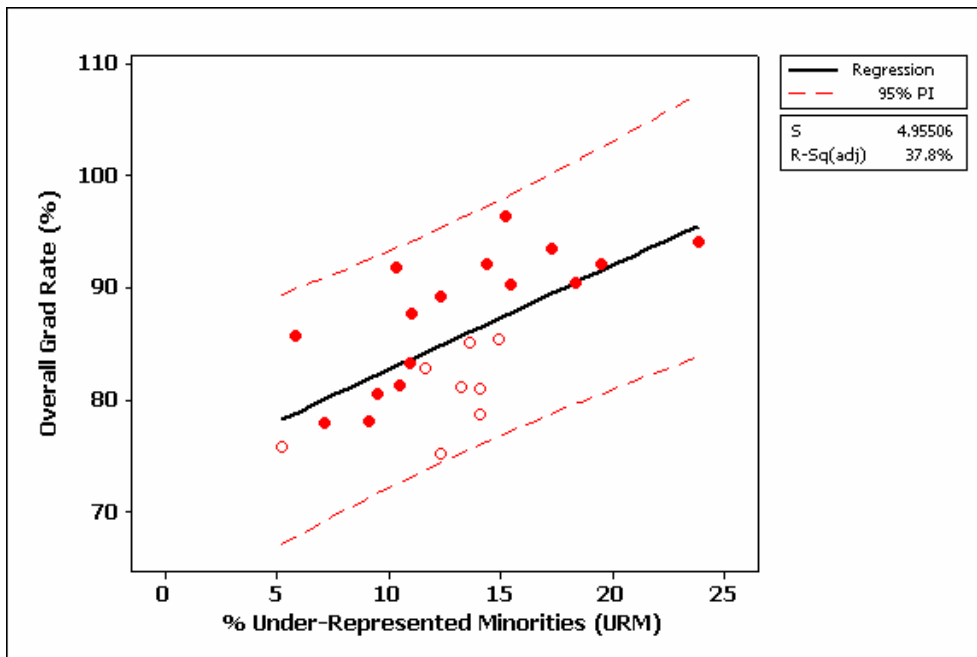
Figure 5: Six-Year Overall Graduation Rate versus Percent of Students who are classified as Under-Represented Minorities (URM)
(Solid circles denote private universities; open circles denote public universities)

When both the median SAT and %URM are included as predictors in a regression for the response to the 6-year overall graduation rate, the adjusted R^2 is 74%. The regression equation is
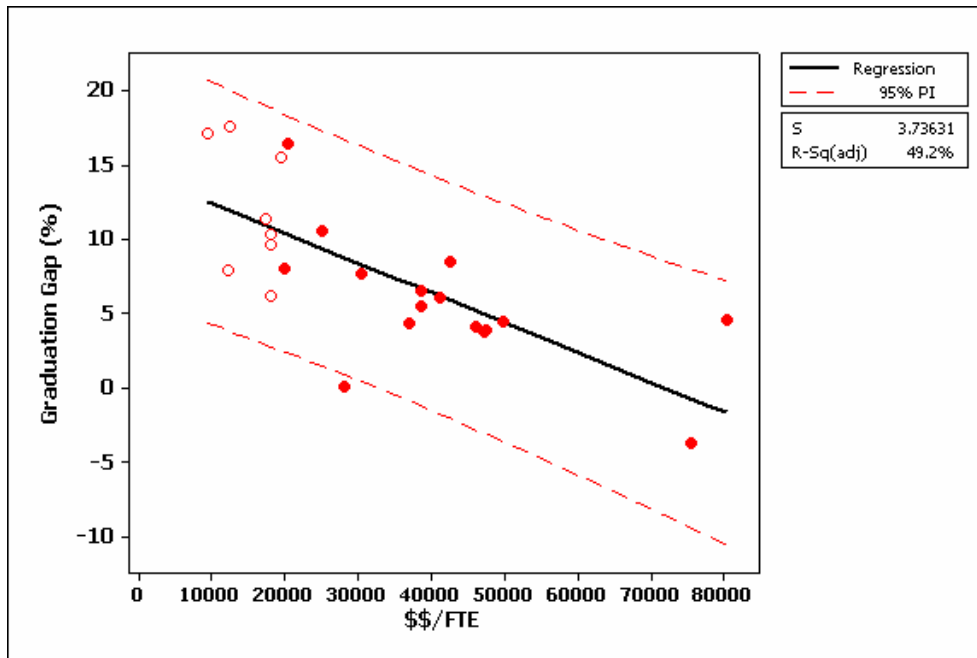
Graduation Rate = 16.9 + 0.0469 Median SAT +   0.473 (%URM)

An analysis of covariance including the type of university (private or public) showed no significant difference between public and private universities when the median SAT and % URM is included in the model.

In addition, the graduation rate was regressed against the student related expenditures per full time student ($$/FTE) and its significance was less.

Graduation Rate Gap

The graduation rate gap is defined as the difference of the 6-year URM graduation rate from the overall 6-year graduation rate (all students). The average graduation rate gap is 7.8%. To consider predictors of the Graduation Rate Gap, a stepwise regression was run. Included were the median SAT, percent of student body that is URM, number of students and percent of students with Pell grants. (See Appendix for data detail.) The student related expenditures/FTE was identified as the most important predictor for graduation gap (p = .000) with a R^2 of 49%. In Figure 6, the graph of the graduation gap versus the student related expenditures/FTE is presented. ***The more money a university invests in instruction, student services and academic support related to instruction, the more successful they are at narrowing the graduation gap.***

Figure 6: Graduation Rate Gap vs. Student related expenditures/FTE
(Solid circles denote private universities; open circles denote public universities)

Clearly, the private universities are more successful with achieving an institutional effort of less graduation rate gap than the public universities. There is a strong relationship between the gap performance and the amount of money available. In fact, one university had a negative gap of 3%.

An Analysis of Covariance including the type of university (private or public) and the Student related expenditure /FTE showed that the type of university was not a significant factor when the covariate Student-Related Expenditure/FTE was taken into account. See Table 2 for the Analysis of Covariance. From this analysis, it can be concluded that the only difference between public and private universities in this study is that the private universities have more money per student and can implement programs that help students. In other words, the only significant trend is the relationship between the Graduation rate gap and the Student related expenditures/FTE.

**Table 2: Analysis of Covariance for Graduation Rate GAP**

```
Source    DF   Seq SS   Adj SS   Adj MS     F      P
$$/FTE     1   325.29   125.01   125.01   8.85   0.007
Type       1    10.56    10.56    10.56   0.75   0.397
Error     21   296.56   296.56    14.12
Total     23   632.41


S = 3.75790    R-Sq = 53.11%    R-Sq(adj) = 48.64%
```

Student Related Expenditures/FTE Stronger Predictor than Median SAT

The correlation between the Graduation Rate Gap and the median SAT is much less. In a regression of these two variables, the R^2 was only 23%. Because of a high correlation between the student related expenditure/FTE and the median SAT, the median SAT is not a significant predictor given that the student Related Expenditure/FTE is a predictor in the regression equation of Graduation Rate Gap.

It was expected that the median SAT would be a stronger predictor than it was. To understand this phenomenon, a graph of the median SAT and Student related expenditures/FTE was generated (Figure 7).   For public universities, the Student related expenditures/FTE is relatively constant over the range of the median SAT; this makes sense because public universities are limited by the total revenue they receive from the state governments. For the private universities, there is a positive correlation between Student related expenditures/FTE and median SAT with more variation. With more private funds and higher tuitions, the private universities can expend more on student instruction and related services that help students succeed in graduation. In addition, the trend is clear that universities having higher student related expenditures can attract students with higher SAT scores. Students admitted to these universities win in two ways: they are better prepared and have more services available to them to help them be successful as students.
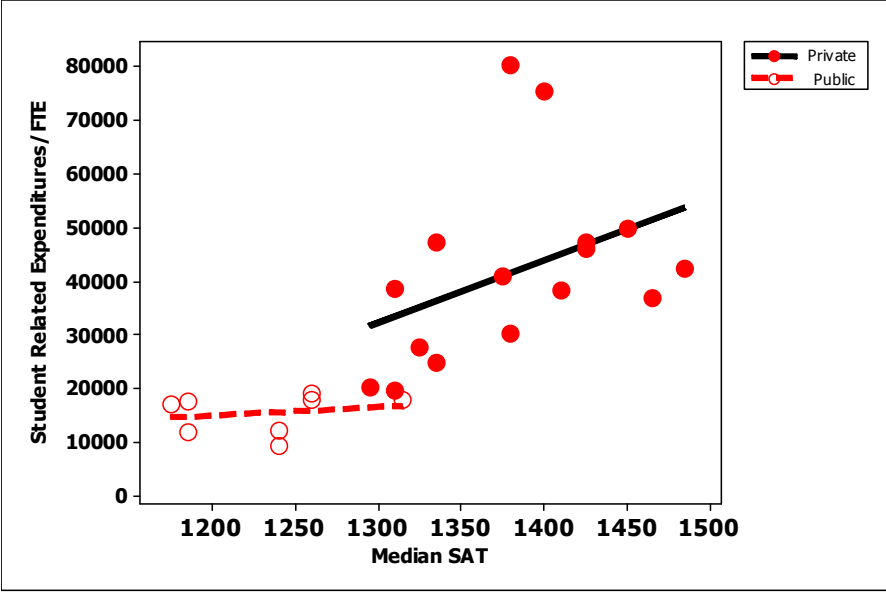
Figure 7: Student related expenditures/FTE versus median SAT

Data Envelopment Analysis

Data Envelopment Analysis (DEA) is a linear programming technique that is often used for benchmarking educational institutions. DEA-Solver software was used to perform a Data Envelopment Analysis on this data. The graduation gap was used as an output variable and the median SAT and Student related expenditures/FTE as input variables. Because the output variable needs to be directionally-scaled as more is better, the gap was subtracted from a constant, in this case 20 (the transformed output is 20- graduation gap).

Both the CCR and BBC – output oriented models were used. For technical reasons related to DEA methodology, the CCR model was used. The DEA analysis indicated that the benchmark frontier included three universities, Texas A&M, Case Western Reserve and Washington University at St. Louis. These universities were efficient with a low gap and corresponding low SAT and low student related expenditures/FTE relative to peer institutions with similar graduation gaps. The values of the variables are shown in Table 3

Table 3: DEA inputs and outputs

University	Gap	20-Gap	Median SAT	Student related expenditures/FTE
Texas A&M	7.9	12.1	1185	12147
Case Western Reserve	0.1	19.9	1325	27986
Washington U. at St. Louis	-3.7	23.7	1400	75512

To visualize this analysis better, Figure 8 exhibits a Minitab 3D graph with all the universities and the (diagonal) line segments represent the frontier planes with the edges of the planes defined by the three benchmark universities identified in Table 3. Clearly, Texas A&M is most efficient because of its low Median SAT and low Student related expenditures/FTE. Even though Washington University has a very high Student related expenditures/FTE, it has a graduation gap of -3.7, causing it to become part of the efficient frontier. Its URM graduation rate was 89% and the overall graduation rate was 93%, clearly a huge success story.
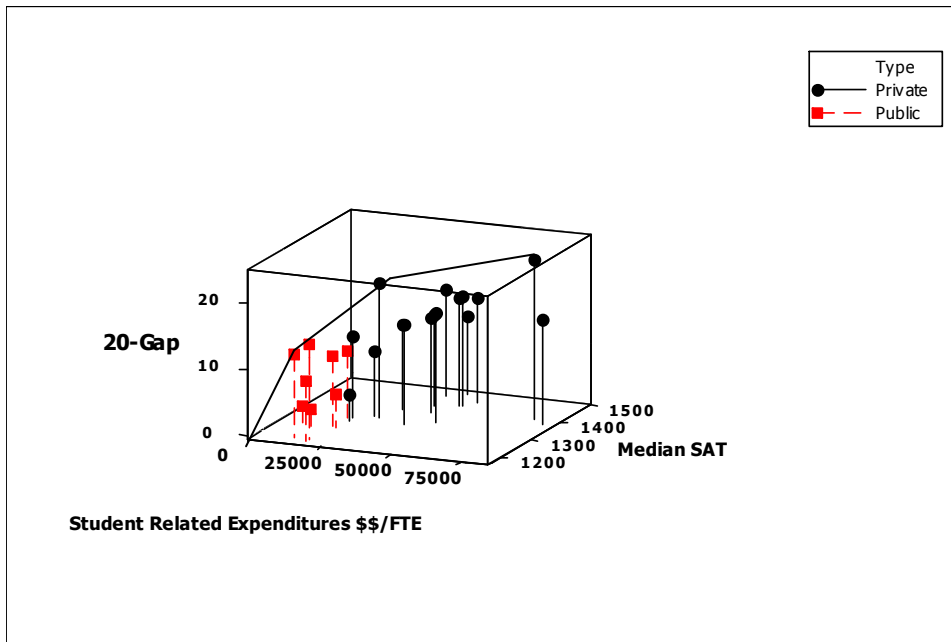
Figure 8: DEA Efficient Frontier: Texas A&M, Case Western Reserve and Washington University.

Comparison of Regression Analysis and DEA

In the regression of the Graduation Gap versus student related expenditures/FTE, both Case Western Reserve and Washington University at St. Louis were identified as outliers while in the DEA analysis; they are part of the efficient benchmark frontier. From Figure 8, the benchmark frontier planes are a distance from most of the other universities, indicating that most universities can improve with a smaller graduation gap. The DEA analysis clearly identifies Case Western Reserve as a university that is being successful as an institution in addressing the graduation gap; yet in the regression analysis this was not clear. The advantage of using the regression analysis initially was that it identified the strong relationships; i.e. which input variables were significant predictors.

Summary and Conclusions

This research looked at the achievement gap in graduation rates for successful large research universities (graduation rates larger than 75%) with a high percent of STEM majors (greater than 25%) The following conclusions can be made from this research:

1. The variation in graduation rate for a specific university tends to improve over time with a decrease in variation. The average improvement in the 6-year graduation rates over six years for this group of universities averages 3.2% with a 95% confidence interval of 2.1 to 4.3%. All but two of the universities improved their graduation rates in six years. The variation in the five year improvement in graduation rates varied from -1% to 10%. This finding is very significant in studies of retention rates of universities.

2. Both the median SAT and percent of the students who are URM are excellent predictors of the overall graduation rate. There was no significant difference in graduation rates between public and private universities when the median SAT and %URM is taken into account.

3. Among the large research universities with high graduation rates, there is a wide variation in the graduation rate gap between the overall student population and the underrepresented-minority population.

4. The student related expenditures/FTE, a measure of money available for instruction and related services at a university, is a good predictor of the graduation rate gap. The relationship between the graduation rate gap and student related expenditures/FTE is particularly strong with an R^2 of 49%. There is no statistically significant difference between private and public universities once the student related expenditures/FTE is taken into account.

5. When the dependent variable is overall graduation rate, the student related expenditures/FTE is a weak predictor compared to median SAT. When the dependent variable is graduation rate gap, the reverse is evident; student related expenditures/FTE is a much stronger predictor than the median SAT.

6. Success stories do exist. Among the best research universities in addressing and being successful in reducing the graduation gap were Texas A&M University, Case Western Reserve and Washington University at St. Louis. It should be noted that Texas A&M is a public university.

7. Regression analysis and Data Envelopment Analysis work well together to more clearly understand the relationship among graduation gap, the median SAT and Student related expenditures/FTE at an institution level.

Recommendation for an ASEE sponsored database on graduation rates of engineering colleges

Currently ASEE does not include graduation rates or first year retention in its engineering college database. The need for solutions to poor retention of engineering students is real. The research in this paper demonstrates the type of analysis that could be achieved if a national database on institutional level statistics related to graduation rates for engineering colleges existed. It is recommended that ASEE aggressively move towards including graduation rates and first year retention in their database of engineering colleges.

References

1. Carey, K., "One Step from the Finish Line", Jan. 2005, Education Trust website, www2.edtrust/org/edtrust, viewed September 5, 2005.

2. Carnegie Foundation website**,** www.carnegiefoundation.org/classification/index.htm

3. Cooper, W., Seiford, L, and Tone, K., (1999) *Data Envelopment Analysis with DEA-Solver Software,* Kluwer Academic Publishers, Boston, MA

4. Education Trust website, www2.edtrust/org/edtrust, viewed September 3, 2005.

5. Tinto, V., (1993), *Leaving College: Rethinking the Causes and Cures of Student Attrition, (2nded.),* The University of Chicago Press, Chicago, Ill

Appendix: Listing of Data

University/DMU	Pub/Priv	Median SAT	Student related expenditures/FTE	Size	%URM	URM Grad	GAP	20-Gap
U Of Wisconsin	Public	1240	12426	27711	5.2	58.2	17.6	2.4
U Of Illinois At Urbana	Public	1240	9456	28472	14.1	63.8	17.2	2.8
Lehigh University	Private	1295	20343	4643	5.8	69.2	16.5	3.5
U Of Michigan	Public	1260	19374	23714	13.6	69.6	15.5	4.5
U of C-Davis	Public	1175	17289	21799	13.2	69.7	11.4	8.6
Tufts University	Private	1335	25133	4831	15.4	79.7	10.6	9.4
U Of C-Berkeley	Public	1315	17966	22363	14.9	75.0	10.4	9.6
U Of C-San Diego	Public	1260	18077	19299	11.6	73.2	9.7	10.3
MIT	Private	1485	42644	4084	19.5	83.6	8.5	11.5
RPI	Private	1310	19917	5173	9.5	72.5	8.0	12.0
Texas A & M U	Public	1185	12147	33901	12.3	67.3	7.9	12.1
Carnegie Mellon U	Private	1380	30364	5312	10.5	73.7	7.7	12.3
Rice University	Private	1410	38509	2842	18.3	83.9	6.6	13.4
U Of C-Irvine	Public	1185	17935	19456	14.1	72.5	6.2	13.8
Cornell University	Private	1375	41216	13638	10.3	85.8	6.1	13.9
U Of Rochester	Private	1310	38637	4433	9.1	72.6	5.5	14.5
Johns Hopkins U	Private	1380	80457	5078	11.0	83.1	4.6	15.4
Stanford U	Private	1450	49952	6646	23.8	89.6	4.5	15.5
Princeton U	Private	1465	36955	4676	15.2	92.1	4.4	15.6
Duke University	Private	1425	46135	6195	17.3	89.5	4.1	15.9
Columbia University	Private	1425	47490	6543	14.4	88.2	3.9	16.1
Vanderbilt University	Private	1335	47348	6248	10.9	79.5	3.8	16.2
Case West.Res U	Private	1325	27986	3405	7.1	77.8	0.1	19.9
Washington U In St Louis	Private	1400	75512	6337	12.3	93.0	-3.7	23.7